

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2003, 16:11:58 ; Search time 9.77143 Seconds
(without alignments)
91.441 Million cell updates/sec

Title: US-09-905-691-2
Perfect score: 19
Sequence: 1 AAAAAAAAAARRARAE 19

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 127863 seqs, 47026705 residues
Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	1	7	36.8	145	1 YV40_DEIRA	Q9ZS4 deinoceus
2	2	7	36.8	193	1 VCO7_ADE04	Q56831 human adeno
3	3	7	36.8	263	1 CMC_BRAJA	P30962 bradyrhizob
4	4	7	36.8	291	1 YN32_MYCTU	Q10515 mycobacteri
5	5	7	36.8	326	1 Y083_CAUCR	P37895 caulobacter
6	6	7	36.8	362	1 CKRA_HUMAN	P46092 homo sapien
7	7	7	36.8	381	1 ARGJ_THETH	P96137 thermus the
8	8	7	36.8	383	1 SUC2_STRCO	Q87840 streptomyce
9	9	7	36.8	387	1 ARGJ_METKA	Q87840 streptomyce
10	10	7	36.8	394	1 FXD3_CHICK	Q87840 streptomyce
11	11	7	36.8	417	1 PROA_MEIRU	P79772 gallus gall
12	12	7	36.8	423	1 CES5_HUMAN	Q86053 meiothermus
13	13	7	36.8	428	1 Y486_MYCLE	Q86053 meiothermus
14	14	7	36.8	444	1 VGLX_HSVBS	P34138 mycobacteri
15	15	7	36.8	448	1 TRNE_XANAC	Q86103 bovine herp
16	16	7	36.8	474	1 TYAE_BORPE	Q86103 bovine herp
17	17	7	36.8	480	1 Y486_MYCTU	Q86103 bovine herp
18	18	7	36.8	499	1 GSHR_PLAF7	Q11522 mycobacteri
19	19	7	36.8	499	1 GSHR_PLAFK	Q11522 mycobacteri
20	20	7	36.8	521	1 EX7L_RHIL0	Q34655 plasmodium
21	21	7	36.8	603	1 U526_HCMVA	Q98703 rhizobium l
22	22	7	36.8	621	1 HEM1_AGABI	P09699 human cytom
23	23	7	36.8	742	1 UL47_HSVBP	Q92403 agaricus bi
24	24	7	36.8	777	1 METE_CAUCR	P30021 bovine herp
25	25	7	36.8	799	1 SYFB_CAUCR	Q9841 caulobacter
26	26	7	36.8	913	1 VGLB_PRVIF	Q9845 caulobacter
27	27	7	36.8	1318	1 VP14_EBV	P08355 pseudorabie
28	28	7	36.8	1394	1 HAP_HAEIN	P03179 epstein-bar
29	29	7	36.8	2390	1 SPCP_HUMAN	P45387 haemophilus
30	30	7	36.8	4523	1 DYHB_HUMAN	O15020 homo sapien
31	31	6	31.6	68	1 RPOZ_NEIMA	Q96dt5 homo sapien
32	32	6	31.6	74	1 Y082_BPF2	Q9jq99 neisseria m
33	33	6	31.6	88	1 RPOZ_HAEIN	Q06424 bacterioph
						P43740 haemophilus

34	6	31.6	89	1	RPOZ_PASMU	Q9cmb2 pasteurella
35	6	31.6	90	1	RPOZ_VIBCH	Q9kmb3 vibrio chol
36	6	31.6	90	1	RPOZ_VIBPA	Q87kb0 vibrio para
37	6	31.6	90	1	RPOZ_VIBVU	Q8ddv5 vibrio vuln
38	6	31.6	91	1	RPOZ_ECOLI	P08374 escherichia
39	6	31.6	91	1	RPOZ_YERPE	Q8zjq3 yersinia pe
40	6	31.6	94	1	R28A_MYCTU	Q10879 mycobacteri
41	6	31.6	108	1	NIFW_RHOSH	Q01182 rhodobacter
42	6	31.6	113	1	DL67_HCMVA	P16747 human cytom
43	6	31.6	117	1	RL18_HAEIN	P43556 haemophilus
44	6	31.6	117	1	VGLJ_HSVBS	P36344 simian herp
45	6	31.6	120	1	PAND_ALCEU	Q9zhi5 alcaligenes

ALIGNMENTS

RESULT 1
YV40_DEIRA
ID YV40_DEIRA STANDARD; PRT; 145 AA.
AC Q9ZS4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein DRB0040.
GN DRB0040.
OS Deinococcus radiodurans.
OG Plasmid MPI.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID-1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RI / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE-20036896; PubMed-10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.N.;
RT *Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.*;
RL Science 286:1571-1577(1999).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AE001826; AAF12646.1;
CC DR PIR: E75622; E75622.
CC TIGR: DRB0040;
CC InterPro: IPR001844; Chaprln_Cpn60.
CC Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 145 AA; 15152 MW; B8ED524495897EC3 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 AARRAR 9
Db 132 AARRAR 138
|||||

RESULT 2
VCO7_ADE04
ID VCO7_ADE04 STANDARD; PRT; 193 AA.

RA Wang W., Soto H., Oldham E.R., Buchanan M.E., Honey B., Catron D.,
 RA Jenkins N., Copeland N.G., Gilbert D.J., Nguyen N., Abrams J.,
 RA Kershenovitch D., Smith K., McClanahan T., Vicari A.P., Zlotnik A.,
 RT Identification of a novel chemokine (CCL28), which binds CCR10
 (GPR2).
 RL J. Biol. Chem. 275:22313-22323(2000).
 RN [5]
 CC LIGAND BINDING.
 RX MEDLINE-20432268; PubMed-10975800;
 RA Pan J., Kunkel E.J., Gossler U., Lazarus N., Langdon P., Broadwell K.,
 RA Vierra M.A., Genovese M.C., Butcher E.C., Soler D.;
 RT A novel chemokine ligand for CCR10 and CCR3 expressed by epithelial
 cells in mucosal tissues.
 RL J. Immunol. 165:2943-2949(2000).
 CC -1- FUNCTION: RECEPTOR FOR CHEMOKINES SCYA27 AND SCYA28. SUBSEQUENTLY
 CC TRANSDUCES A SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS
 CC LEVEL AND STIMULATES CHEMOTAXIS IN A PRE-B CELL LINE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT TESTIS,
 CC SMALL INTESTINE, FETAL LUNG, FETAL KIDNEY. WEAKER EXPRESSION WAS
 CC OBSERVED IN MANY OTHER ADULT TISSUES INCLUDING SPLEEN, THYMUS,
 CC LYMPH NODE, PEYER'S PATCHES, COLON, HEART, OVARY, PERIPHERAL BLOOD
 CC LYMPHOCYTES, THYROID AND SPINAL CORD. ALSO EXPRESSED BY
 CC MELANOCYTES, DERMAL FIBROBLASTS, DERMAL MICROVASCULAR ENDOTHELIAL
 CC CELLS. ALSO DETECTED IN T CELLS AND IN SKIN-DERIVED LANGERHANS
 CC CELLS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF215981; AAF63709.1; -;
 DR EMBL; AF208237; AAF72871.1; -;
 DR EMBL; U13667; AAA64593.1; -;
 DR Genew; HGNC:4474; GPR2.
 DR MIM; 600240; -;
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. .; TAS.
 DR InterPro; IPR005382; CC.chemkinel0.
 DR Genew; HGNC:4474; GPR2.
 DR Pfam; PF00001; 7tm1.1;
 DR PRINTS; PR01557; CHEMOKINER10.
 DR PRINTS; PR00237; GPCRHOODPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 53 68 1 (POTENTIAL).
 FT DOMAIN 59 78 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 79 99 2 (POTENTIAL).
 FT DOMAIN 100 114 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 115 136 3 (POTENTIAL).
 FT DOMAIN 137 159 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 160 179 4 (POTENTIAL).
 FT DOMAIN 180 203 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 204 225 5 (POTENTIAL).
 FT DOMAIN 226 247 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 248 269 6 (POTENTIAL).
 FT DOMAIN 270 290 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 291 313 7 (POTENTIAL).
 FT DOMAIN 314 362 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 113 191 BY SIMILARITY.
 FT CONFLICT 51 51 V -> L (IN REF. 3).
 FT CONFLICT 235 235 MISSING (IN REF. 2).
 FT CONFLICT 330 330 S -> C (IN REF. 2).
 SQ SEQUENCE 362 AA; 38399 MW; 69DF12B639AE99A CRC64;

Query Match

36.8%; Score 7; DB 1; Length 362;

Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 AARRAAR 9
 DB 68 AARRAAR 74
 RESULT 7
 ARGJ_TRETH
 ID ARGJ_TRETH STANDARD; PRT; 381 AA.
 AC P96137;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glutamate N-acetyltransferase (EC 2.3.1.35) (Ornithine
 DE acetyltransferase) (Ornithine transacetylase) (OATase) [Contains:
 DE Glutamate N-acetyltransferase alpha chain; Glutamate N-
 DE acetyltransferase beta chain].
 GN ARGJ.
 OS Thermus thermophilus.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermaceae;
 OC Thermus.
 OX NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB27;
 RX MEDLINE-98154436; PubMed-9493385;
 RA Baetens M., Legrain C., Boyen A., Glansdorff N.;
 RT 'Genes and enzymes of the acetyl cycle of arginine biosynthesis in the
 RT extreme thermophilic bacterium Thermus thermophilus HB27.';
 RL Microbiology 144:479-492(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB27;
 RA Sanchez R., Roovers M., Glansdorff N.;
 RT 'Organisation of arginine biosynthetic genes in Thermus
 RT thermophilus.';
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + L-glutamate -> L-
 CC ornithine + N-acetyl-L-glutamate.
 CC -1- PATHWAY: Arginine biosynthesis; fifth step.
 CC -1- SUBUNIT: Heterotetramer of two alpha and two beta chains (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- MISCELLANEOUS: Some bacteria possess a monofunctional argJ, i.e.,
 CC capable of catalyzing only the fifth step of the arginine
 CC biosynthetic pathway.
 CC -1- SIMILARITY: Belongs to the argJ family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y10525; CAA71551.1; -;
 DR EMBL; Y18353; CAA7143.1; -;
 DR HAMAP; MF_01106; -; 1.
 DR InterPro; IPR002813; ArgJ.
 DR Pfam; PF01960; ArgJ; 1.
 DR PRODOM; PD004193; ArgJ; 1.
 DR TIGRFAMs; TIGR00120; ArgJ; 1.
 KW Arginine biosynthesis; Transferase; Acyltransferase.
 FT CHAIN 1 175 GLUTAMATE N-ACETYLTRANSFERASE ALPHA CHAIN
 FT CHAIN 176 381 (BY SIMILARITY).
 FT CHAIN 188 189 GLUTAMATE N-ACETYLTRANSFERASE BETA CHAIN
 FT SITE 381 AA; 40318 MW; 1605CSB17B7B05A8 CRC64;
 SQ SEQUENCE 381 AA; 40318 MW; 1605CSB17B7B05A8 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 381;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ARAARA 10
DB 274 ARAARA 280

RESULT 8
SUC2_STRCO STANDARD; PRT; 383 AA.
AC O87840;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA synthetase beta chain 2 (EC 6.2.1.5) (SCS-beta 2).
GN SUCC OR SC06585 OR SC8A6.06.

OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-A3(2) / M145;
RX MEDLINE-21996410; PubMed-12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Kiese H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Crohn A., Fraser A., Goble A., Hidaigo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kiese T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).
RL Nature 417:141-147(2002).

CC -1- CATALYTIC ACTIVITY: ATP + succinate + CoA = ADP + succinyl-CoA +
phosphate.
CC -1- PATHWAY: Tricarboxylic acid cycle.
CC -1- SUBUNIT: Composed of an alpha chain and a beta chain (By
similarity).
CC -1- SUBUNIT: Belongs to the succinate/malate CoA ligase beta
subunit family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; AL939128; CAA19778.1; -;
CC PIR; T35773; T35773.
CC HAMAP; P07460; 18CU.
CC HSSP; P07460; 18CU.
CC InterPro; IPR003135; ATP-grasp.
CC InterPro; IPR005809; CoA_lig_beta.
CC InterPro; IPR005811; CoA_ligase.
CC Pfam; PF02222; ATP-grasp; 1.
CC Pfam; PF00549; Ligase-CoA; 1.
CC TIGRFAMs; TIGR01016; sucCoABeta; 1.
CC PROSITE; PS01217; SUCCINYL_COA_LIG_3; 1.
KW Ligase; Tricarboxylic acid cycle; Complete proteome.
SQ SEQUENCE 383 AA; 40161 MW; 5A3D05C0FD1DD03 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 383;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 AARRAR 9

DB 365 AARRAR 371

RESULT 9
ARGJ_METKA STANDARD; PRT; 387 AA.
AC Q8TX15;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Arginine biosynthesis bifunctional protein argJ (includes: Glutamate
N-acetyltransferase (EC 2.3.1.35) (Ornithine acetyltransferase)
(Ornithine transacetylase) (OATase); Amino-acid acetyltransferase
(EC 2.3.1.1) (N-acetylglutamate synthase) (AGS)) [Contains: Arginine
biosynthesis bifunctional protein argJ alpha chain; Arginine
biosynthesis bifunctional protein argJ beta chain].
GN ARGJ OR MKQ865.

OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-AV19 / DSM 6324 / JCM 9639;
RX MEDLINE-21927647; PubMed-11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natile D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozavkin S.A.;
RT The complete genome of hyperthermophile Methanopyrus kandleri AV19
and monophyly of archaeal methanogens.
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
CC -1- FUNCTION: Catalyzes two activities which are involved in the
cyclic version of arginine biosynthesis: the synthesis of
acetylglutamate from glutamate and acetyl-CoA, and of ornithine by
transacetylation between acetylornithine and glutamate (By
similarity).
CC -1- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + L-glutamate = L-
ornithine + N-acetyl-L-glutamate.
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + L-glutamate = CoA + N-acetyl-L-
glutamate.
CC -1- PATHWAY: Arginine biosynthesis; first step.
CC -1- PATHWAY: Arginine biosynthesis; fifth step.
CC -1- SUBUNIT: Heterotetramer of two alpha and two beta chains (By
similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- MISCELLANEOUS: Some bacteria possess a monofunctional argJ, i.e.,
capable of catalyzing only the fifth step of the arginine
biosynthetic pathway.
CC -1- SIMILARITY: Belongs to the argJ family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; AE010376; AA02078.1; -;
CC HAMAP; MF_01106; atypical; 1.
CC InterPro; IPR002813; ArgJ.
CC Pfam; PF01960; ArgJ; 1.
CC ProDom; PD004193; ArgJ; 1.
CC TIGRFAMs; TIGR00120; ArgJ; 1.
CC Arginine biosynthesis; Multifunctional enzyme; Transferase;
KW Acyltransferase; Complete proteome.
FT CHAIN 1 172 ARGinine BIOSYNTHESIS BIFUNCTIONAL
FT CHAIN 173 387 PROTEIN ARGJ ALPHA CHAIN (BY SIMILARITY).
FT CHAIN 387 PROTEIN ARGJ BETA CHAIN (BY SIMILARITY).
SQ SEQUENCE 387 AA; 41840 MW; ED311EC1F47D56DD CRC64;

Query Match 36.8%; Score 7; DB 1; Length 387;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ARRAARA 10
 DB 276 ARRAARA 282

RESULT 10

FXD3_CHICK STANDARD; PRT; 394 AA.
 AC P79772;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Forkhead box protein D3 (HNF3/FH transcription factor, genesis) (Winged
 helix protein CWH-3).
 GN FOXD3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=97141794; PubMed=8988052;
 RA Freyaldenhoven B.S., Freyaldenhoven M.P., Iacovoni J.S., Vogt P.K.;
 RT Aberrant cell growth induced by avian winged helix proteins.";
 RL Cancer Res. 57:123-129(1997).
 CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 1 fork-head domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC

EMBL; U37274; AAC60066.1; -
 DR HSSP; Q63245; 2HFH.
 DR TRANSFAC; T02495; -
 DR InterPro; IPR001766; TF_Fork_head.
 DR Pfam; PF00250; Fork_head; 1.
 DR PRINTS; PR00053; FORKHEAD.
 DR ProDom; PD000425; TF_Fork_head; 1.
 DR SMART; SM00339; FH; 1.
 DR PROSITE; PS00657; FORK_HEAD_1; 1.
 DR PROSITE; PS00658; FORK_HEAD_2; 1.
 DR PROSITE; PS50039; FORK_HEAD_3; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation.
 FT DOMAIN 67 70 POLY-ALA.
 FT DOMAIN 80 91 POLY-GLY.
 FT DOMAIN 100 106 POLY-ALA.
 FT DNA_BIND 117 211 FORK-HEAD.
 SQ SEQUENCE 394 AA; 40995 MW; 324A4B36B9E31899 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 394;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAARRAA 8
 DB 62 RAARRAA 68

RESULT 11

PROA_MEIRU

PROA_MEIRU STANDARD; PRT; 417 AA.
 AC Q86033;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-
 semialdehyde dehydrogenase) (Glutamyl-gamma-semialdehyde
 dehydrogenase) (GSA dehydrogenase).
 GN PROA.
 OS Meiothermus ruber.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 OC Meiothermus.
 OX NCBI_TaxID=277;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=40;
 RA Yaklichkin S.Y., Zimina M.S., Yurchenko Y.V., Hromov I.S.,
 RT "Molecular cloning and sequence analysis of the proA gene from
 thermophilic eubacterium Thermus ruber.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE NADPH DEPENDENT REDUCTION OF L-GAMMA-
 GLUTAMYL 5-PHOSPHATE INTO L-GLUTAMATE 5-SEMIALDEHYDE AND
 PHOSPHATE. THE PRODUCT SPONTANEOUSLY UNDERGOES CYCLIZATION TO-FORM
 1-PYRROLINE-5-CARBOXYLATE.
 CC -1- CATALYTIC ACTIVITY: L-glutamate 5-semialdehyde + phosphate +
 NADP(+) -> L-gamma-glutamyl 5-phosphate + NADPH.
 CC -1- PATHWAY: Proline biosynthesis; second step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE GAMMA-GLUTAMYL PHOSPHATE REDUCTASE
 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF082661; AAC72811.1; -
 DR HAMAP; MF_00412; -; 1.
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR InterPro; IPR000965; Gglut_pp_reduct.
 DR Pfam; PF00171; algedh; 1.
 DR TIGRfam; TIGR00407; proA; 1.
 DR PROSITE; PS01223; PROA; 1.
 DR Oxidoreductase; Proline biosynthesis; NADP.
 KW SEQUENCE 417 AA; 44985 MW; A280A8A7E9C92268 CRC64;
 SQ
 Query Match 36.8%; Score 7; DB 1; Length 417;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ARAAAR 13
 DB 13 ARAAAR 19

RESULT 12
 CESS_HUMAN STANDARD; PRT; 423 AA.
 ID CESS_HUMAN
 AC Q9BXW7; Q9BXW8; Q9NWA8; Q9NX41;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cat eye syndrome critical region protein 5 precursor.
 GN CECR5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

```

ENBL; AF273271; AAK19152.1; -
ENBL; AF273270; AAK19151.1; -
ENBL; AK001034; BAA91475.1; -
ENBL; AK000461; BAA91180.1; -
ENBL; BC042540; AAH42540.1; -
Genew; HGNC:1843; CECR5.
DR InterPro; IPR006353; HAD_CECR5.
DR InterPro; IPR006357; HAD_SF_IIA.
DR TIGRFAMs; TIGR01456; CECR5; 1.
DR TIGRFAMs; TIGR01460; HAD-SF-III; 1.
FT SIGNAL; 1 23 POTENTIAL:
FT CHAIN 24 423 CAT EYE SYNDROME CRITICAL REGION PROTEIN
FT FT VARSPLIC 1 41 5.
FT FT --> MYAFVLPSPS (in isoform 1).
FT FT /Tid=VSP_003840.
FT FT D -> N (IN REF. 2; BAA91180).
FT FT CONFLICT y 54 54
FT FT CONFLICT 86 86 V -> F (IN REF. 2; BAA91475).
FT FT CONFLICT 297 297 E -> G (IN REF. 2; BAA91475).
SQ SEQUENCE 423 AA; 46321 MW; C4D9208AB8B8CCE CRC64;

Query Match 36.8%; Score 7; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RAARAAA 12
DD 19 RAARAAA 25
|||||

RESULT 13
ID Y486_MYCLE STANDARD; PRT; 428 AA.
AC P54138; Q9CB50;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE DE Hypothetical protein ML2443.
GN ML2443 OR U2169F OR B2168_C2_201.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TAXID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Elgmelter K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Dutboy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC CC -! SIMILARITY: TO M.TUBERCULOSIS RV0486.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to licenses@isb-sib.ch).
-----

```

```

DR EMBL; U00018; AAA17228.1; ALT_INIT.
DR EMBL; AL583925; CAC31960.1; -.
DR PIR; H87214; H87214.
DR Lepitoma; ML2443; -.
DR InterPro; IPR001296; Glyco_trans_1.
DR Pfam; PF00534; Glycos_trans_1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 428 AA; 45291 MW; A14F9F0187E3587C CRC64;

Query Match 36.8%; Score 7; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RRAARAA 11
DB 204 RRAARAA 210
|||||

RESULT 14
VGLX_HSVBS STANDARD; PRT; 444 AA.
AC Q08103;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein G precursor.
OS Bovine herpesvirus type 1.2 (strain ST).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=45407;
RN [1]

SEQUENCE FROM N.A.
RX MEDLINE=94167875; PubMed=8122370;
RA Leung-Tack P., Audonnet J.F., Riviere M.;
RT "The complete DNA sequence and the genetic organization of the short
RL unique region (US) of the bovine herpesvirus type 1 (ST strain).";
Virology 199;409-421(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; Z23068; CA80603.1; -.
DR PIR; S35783; S35783.
DR InterPro; IPR003363; Herpes_gg.
DR Pfam; PF02400; Herpes_gg; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 444 GLYCOPROTEIN GX.
FT TRANSMEM 390 414 POTENTIAL.
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 444 AA; 46708 MW; 0145942AA35B05CB CRC64;

Query Match 36.8%; Score 7; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AARRARA 17
DB 422 AARRARA 428
|||||

RESULT 15
TRME_XANAC STANDARD; PRT; 448 AA.
AC O8PEH9;
DT 28-FEB-2003 (Rel. 41, Created)

```

```

DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE tRNA modification GTPase trme.
GN TRME OR THDF OR XAC4370.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]

SEQUENCE FROM N.A.
RP STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergro F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417;459-463(2002).

CC -!- FUNCTION: Exhibits a very high intrinsic GTPase hydrolysis rate.
CC Involved in the biosynthesis of the hypermodified nucleoside 5-
CC methylaminomethyl-2-thiouridine, which is found in the wobble
CC position of some tRNAs (By similarity).
CC -!- SIMILARITY: Belongs to the era/trme family of GTP-binding
CC proteins. Trme subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; AE012091; AM39200.1; -.
DR HAMAP; MF_00379; -.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR006073; GTP_OBG.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR004520; ThdF.
DR PRINTS; PR00326; GTP_OBG.
DR TIGRFAMs; TIGR00650; MG442; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR TIGRFAMs; TIGR00450; thdF; 1.
KW tRNA processing; GTP-binding; Complete proteome.
FT NP_BIND 225 232 GTP (POTENTIAL).
FT NP_BIND 272 276 GTP (POTENTIAL).
FT NP_BIND 332 335 GTP (POTENTIAL).
SQ SEQUENCE 448 AA; 47543 MW; B6B43D163D92E3F4 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 448;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAARRA 7
DB 143 ARAARRA 149
|||||

Search completed: August 9, 2003, 16:29:50
Job time : 11.7714 secs

```